



This is one of an introductory overview series of current research, hypotheses, and development methodologies leading to the proposal of global standards for conversational technologies by the Open Voice Network openvoicenetwork.org, an open-source community of The Linux Foundation.

ETHICAL GUIDELINES FOR VOICE EXPERIENCES

A CASE FOR INCLUSIVITY AND TRUSTWORTHINESS

Executive Summary

This report of the Open Voice Network's Ethical Use Task Force was created for enterprises and organizations currently leveraging or considering voice technology. It explores pressing research questions relevant to human interaction with voice technology, such as:

- What are voice-specific issues of ethical and moral concern, and how do we define them?
- What are the values to be promoted for ethical interaction with voice technology?
- What are the rights to be respected for successful, ethical voice experiences?
- What preventative measures can we take to protect people from voice-specific harm?
- What can we do to cover gaps in voice that are not yet fully covered by existing laws and guidelines?
- What can we learn from past voice-specific harmful incidents and vulnerabilities, and how can we prevent them from happening in the future?

The paper outlines a voice-specific ethical framework based on fundamental ethical principles of **Compliance, Transparency, Inclusivity, Privacy Protection, Accountability, and Sustainability** throughout the design process to avoid potential harms such as bias, discrimination, and exclusion for humans interacting with voice technology. Next steps discussed include industry collaboration and assistance integrating an ethical framework .

For the latest, detailed information on the work of the Open Voice Network, please visit our website at openvoicenetwork.org and the Open Voice Network GitHub Repository at github.com/open-voice-network/docs .

Introduction

Who These Guidelines Are For and Why They Matter

These ethical guidelines for voice experiences are for enterprises and organizations currently leveraging voice technology or considering it in the future. They were created by the Open Voice Network's Ethical Use Task Force community to protect individuals who engage with voice technology and could be adversely affected by its use. Our members come at voice tech from every angle, from voice tech designers and practitioners to audio experts, community organizers, voice actors, academics, entrepreneurs, and voice tech users. We believe voice has great value for people.

There is much at stake. Interactive technologies are becoming increasingly voice enabled, creating more opportunities to give people access to the digital world. In particular, marginalized or underserved groups can integrate voice technology into their daily activities without depending on others for help. Because of its inclusive nature, voice can have a real impact and bring tangible improvements to people's lives.

Although there are numerous benefits to using voice, there are currently inherent risks. Voice technology can be used to identify, analyze, and deduce sensitive personal information from the sound of an individual's voice and ambient noise. Misuse of personal voice data, poor data security management, and training with limited datasets can result in personal data leakage, personal data theft, and unfair outcomes. Poor quality training data, limited in size and scope, can lead to frustrating experiences, exclusion, bias, and systemic discrimination.

Our goal is to give everyone an equal opportunity to use voice technology without fearing adverse effects. The first step are these guidelines for optimizing the benefits of voice experiences while reducing risk and avoiding harm from the use of personal voice data. **Increasing the public's ability to trust using voice technology is good for industry, too, as public awareness of privacy and other concerns is increasing.**

About Our Approach

Over the course of 1.5 years, the Ethical Use Task Force met weekly to review, discuss, analyze, and interpret reputable sources of information related to voice and ethics, attended ethics-related events, and brought in experts from various fields to share their knowledge and resources about ethics. This document reflects our discussions, research, and analysis of information brought up in our meetings, the events we attended, and our individual research.

In this document, we outline a voice-specific ethical framework designed to reassure users of the full respect of their rights built into voice use cases. Our ethical principles reflect a belief in a partnership between humans and technology, where technology enhances human capabilities rather than replaces them and contributes to the well-being of individuals, society, and the environment. Additional guidance will be released once there is more clarity on the ethical ramifications of voice technology in Extended Reality, or "real-time interactive 3D" as the "metaverse" develops. To learn about other compelling issues we are working on, see our upcoming reports "Synthetic Voice For Content Owners and Creators" and "Privacy Principles and Capabilities Unique to Voice". Please visit our website at <https://openvoicenetwork.org>. You can contact us on that site, too.

Table of Contents

The Open Voice Network Working Definition of Personal Voice Data	4
Voice Data Analysis	4
Ethical Framework for Trustworthy Voice Experiences	7
Governance/Compliance	8
Transparency	9
Privacy Protection	12
Inclusivity	16
Accountability	20
Sustainability	21
Call to Action	25
About the Open Voice Network	26
About The Linux Foundation	27
Acknowledgements	28
Reference List	28
Licensing and Attribution	33

The Open Voice Network Working Definition of Personal Voice Data

- Direct or inferred information that can identify an individual through the sound of their voice such as a voiceprint, emotional state, and health status
- Personal information spoken out loud by an individual
- Ambient noise such as street noise or people talking in the background
- Personal information linked to an individual's voice such as account numbers, names, contact information, contacts, and device information

Voice Data Analysis

The human voice is composed of distinctive sonic characteristics unique to each individual, making it a valuable tool for enterprises and organizations. These characteristics can be processed and analyzed for various purposes such as:

- Determining and verifying identity
- Assessing emotional state
- Evaluating people for jobs, loans, or other opportunities
- Detecting diseases (Chen, 2019)

The analysis of the voice focuses on the sound used to produce it, defined by Abercrombie (1967) as "those characteristics which are present more or less all the time that a person is talking" (as cited in Laver, 1980, p. 1). These include tone, speed, duration, volume, and pitch.

Voice technology can identify, analyze, and deduce sensitive personal information related to an individual's identity, physical characteristics, physical health, mental health, socioeconomic status, location, emotions, intent, and more.

If not properly managed and secured, this information can be used to make decisions and recommendations regarding people's lives that could potentially cause irreparable harm. Prison convictions have been made based on voice analysis, only to be overturned years later because of contradictory DNA evidence (Claypoole, 2021). These types of decisions and outcomes can be prevented and mitigated if the proper measures and guidelines are in place.

Table 1

Data that can be processed and analyzed from the sound of a voice

Who-What-How	Category	Attribute Used	Example	Legal Definition
Who You Are	Identification	Name and related identity data	Annika	biometric
	Demography	educational level, social class	university degree	
	Ethnicity	language, dialect	German	
	Personality Type	extroversion - introversion	introvert	
	Physical Characteristic	gender identity	female	
	Physical Characteristic	age bracket	20's - 30's	
	Physical Characteristic	height, weight	tall, thin	
	Physical Characteristic	upper body strength	moderate	
Who You Are, Where You Are	Demography	cultural region	Bavarian	
Where You Are	Geography	location - ambient sound	Würzburg	
What You Want	Intent	desired knowledge or outcome	buy bread, milk, beer	
What You Are	Sentiment	emotion	high stress levels; angry	
	Trustworthiness	Intent, believability	believable	
How You Are - Leading Indicators	Physical Condition	intoxication	negative	
	Physical Health	Parkinson's, Alzheimer's	negative	biomarker
	Physical Health	fertility	negative	biomarker
	Physical Health	Infectious respiratory disease	positive	biomarker
	Mental Health	Schizophrenia	negative	biomarker

Note. A table describing how voice technology can identify, analyze, and deduce sensitive personal information related to an individual's identity, physical characteristics, physical health, mental health, socioeconomic status, location, emotions, intent, and more. (This list is illustrative; it is not comprehensive.) References can be found at <https://openvoicenetwork.org>. By Open Voice Network, 2022.

Voice's characteristic as an identifier raises several concerns regarding consent, data usage, data ownership, data security, and data accuracy. The collection and use of biometric data are already regulated by the European General Data Protection Regulation (GDPR) and a few U.S. state laws, such as the Illinois Biometric Information Privacy Act (BIPA). However, many ethical issues surrounding voice analysis run into broader questions, such as:

- When is it okay to use the information?

- How much information is acceptable to use?
- Is the information accurate?
- What constitutes discrimination or bias?

While anti-discrimination laws do exist, they do not yet fully cover discrimination that can occur specifically from the sound of the voice.

Voice analysis can generate a complete personality profile and correlate the profile to sensitive categories such as race, ethnicity, or gender. This can lead to bias and discrimination. For example, as more companies use voice technology to help facilitate hiring/interview decisions, a candidate could be denied a job due to algorithms favoring specific attributes.

According to Burns (2021), "...[S]oftware like AON's vidAssess scans a recorded video interview for spoken words and phrases that align with the competencies you are looking for in an employee and scores them for an at-a-glance overview," although "early signals suggest certain accents and tones are being prioritized" over others (paras. 16 & 20).

In a study conducted by Dr. John Baugh (2002), landlords were more likely to deny housing opportunities to prospective tenants of color based solely on the sound of their voice, regardless of their qualifications. Baugh found that this type of "linguistic profiling" occurred when searching for employment, service providers, and loans as well (Rice, 2002, para. 2).

In the medical field, specific features and signals in the sound of the voice known as biomarkers can potentially be analyzed to detect, treat, monitor, and prevent cardiovascular diseases such as:

- Coronary artery disease
- Neurological disorders, such as Amyotrophic lateral sclerosis (ALS), Parkinson's, and Alzheimer's
- Psychiatric disorders, such as schizophrenia, bipolar disorder, post-traumatic stress disorder, and depression (Fagherazzi et al., 2021)

Promising research has also been done on the use of voice biomarkers to detect Covid-19 in individuals through their voice and respiratory sounds, such as coughing and breathing (Despotovic et al., 2021). One study that detected Covid-19 through cough patterns achieved 88.52% accuracy, 88.75% sensitivity, and 90.87% specificity (Despotovic et al., 2021).

Voice analysis shows encouraging results in detecting Parkinson's disease, in particular. In a study conducted to determine if algorithms could accurately detect Parkinson's disease in individuals, 6,500 sound samples were collected from a group of around 6,000 people (Wroge et al., 2018). Algorithms were then used to determine which of those people had Parkinson's and which people did not, with a peak accuracy rate of 85% (Wroge et al., 2018). Although this research is promising, many more voice samples are needed to increase the accuracy and precision of diagnoses (Wanucha, 2019).

As with other biomarkers, voice biomarker analysis comes with concerns of accuracy, reliability, specificity, sensitivity, variability, bias, privacy, and interpretation, especially when it comes to something as serious as diagnosing diseases (Mayeux, 2004). To date, no digital voice biomarker technology has been approved by the United States Food and Drug Administration (FDA), so it is

important not to rely solely on voice analysis when making important medical decisions (Wroge et al., 2018).

Extra caution should also be used when relying on Speech Emotion Recognition (SER) and sentiment analysis to personalize experiences or market products to people. In voice experiences, both spoken words and the sound of the voice can be analyzed to determine an individual's emotional state. Sentiment analysis determines how people are feeling based on the words they use, and Speech Emotion Recognition determines how people are feeling based on the sound of their voice (Ingale & Chaudhari, 2012). Emotional states such as happiness, sadness, fear, and anger can potentially be detected as well as physiological states, such as intoxication and hunger (Maghilnan & Kumar, 2017).

The information that can be inferred from these emotions is highly sensitive and personal and using the information could unintentionally make situations worse in some cases. Emotions are tricky to interpret, as they are often temporary and can easily be misconstrued for many reasons, including cultural differences.

While the intention of using voice analysis might be to improve consumer experiences and create mutually beneficial outcomes, the outcomes that occur are not always desirable or accurate. Decisions should not be made based solely on voice data analysis, and ethical standards and guidelines should be implemented to mitigate harm from the use of personal voice data.

Ethical Framework for Trustworthy Voice Experiences

Figure 1

Open Voice Network ethical principles for voice experiences



Note. Overarching principle of governance and compliance encompassing the other ethical principles of transparency, privacy protection, inclusivity, accountability, and sustainability to achieve trustworthy voice experiences. By Open Voice Network, 2022.

Governance/Compliance

Overarching compliance with all of the ethical principles mentioned, standards, guidelines, and existing laws and regulations will be crucial to achieving trust in voice experiences.

We acknowledge the widely used and accepted description of trust in technology and automated systems by Lee and See (2004), which states that technology “will help achieve an individual’s goals in a situation characterized by uncertainty and vulnerability” (p. 51). While the authors’ definition works generally, the addition of emotion and context in voice technology means that people need additional reassurance beyond the fulfillment of their goals to use it confidently on a regular basis.

While achieving an expected outcome is crucial to gaining trust, it is equally— if not more important— to:

- Uphold and comply with the highest standards and the law
- Maintain open and transparent communication
- Minimize the amount of personal voice data collected
- Continuously monitor and audit decisions and outcomes regarding personal voice data

This will help reduce any uncertainty and vulnerability people might experience before, during, and after any voice interaction.

Follow the law

There are different laws and regulations for voice experiences depending on your location, particularly for privacy, so make sure you are thoroughly informed of [existing laws and regulations](#) wherever people will potentially use your voice experience.

West (2018) says, “Individuals want companies to take meaningful action to protect them from unfairness, bias, poor accountability, inadequate privacy protection, and a lack of transparency. If those steps fail, legislation will become the likely recourse” (sec. 9, para. 2).

Use standards, guidelines, and best practices when developing voice technologies and experiences

Voice-specific laws and regulations are still evolving as voice technology becomes more advanced. Implementing standards, guidelines, and best practices will increase your chances of being compliant with existing laws and regulations and prevent potential compliance issues from occurring in the future.

Transparency

Open and transparent communication is key to building a foundation of trust. Any decisions regarding personal voice data should be disclosed and presented in an easily accessible, understandable, and explainable format.

Should the intent and purpose of collecting personal voice data become vague or nonexistent at any point, trust will erode rapidly. Lack of transparency in the past has led to backlash against major tech companies for sending audio recordings from voice assistants to contractors for review without anyone's knowledge (Barrett, 2019). People are vulnerable every time they interact with voice assistance, so transparency will encourage them to trust that no harm will come from using voice technology.

Be clear about voice data ownership

Living human beings own their voice data unless they explicitly consent to the transfer of ownership. Transfer of ownership should be clearly and explicitly stated in the terms and conditions, and all entities assuming ownership should be listed.

There cannot be any ambiguity when it comes to the ownership of voice data for the protection of people interacting with the voice experience and the organizations and enterprises designing and developing voice experiences. If ownership is transferred, people should clearly understand who owns what voice data and why.

These guidelines address voice data ownership only for living human beings. Future research will be done to address the impact of voice data ownership for the deceased.

Integrate an "ethical black box" in the design

UNI Global Union (2017) suggests one way to enhance transparency is to integrate and implement an "ethical black box," similar to one found in airplanes (p. 7). This black box should be able to record all decisions made and explain the reasoning behind all decisions. In robots, this applies to movement and sensory data as well. The read-out of the decisions and reasoning behind them should be easily understandable, concise, and simple.

Amazon has already started doing this with Alexa. People can ask Alexa, "Why did you do that?" and Alexa will explain why a specific action was taken (Gershgorn, 2019). This technology is still evolving, but understanding what happened and why certain decisions were made will improve the overall experience and strengthen the relationship of trust in technology.

Disclose the type of voice analysis used and the purpose of using it

Make sure people are informed of the type of voice analysis being used, the purpose of using that type of voice analysis, and how the voice data collection and analysis work.

Examples of voice analysis include:

- Voice biometrics for identification and authentication
- Speech Emotion Recognition to detect emotions
- Voice biomarkers for medical purposes.

Disclose the type of personal voice data being collected and how it is being used

During voice interactions, people reveal more personal information about themselves through their voice than they might realize, even for seemingly low-risk interactions such as asking for the weather or playing music. From only one interaction, no matter how small, platform vendors can choose to collect data that includes:

- personal information, including your name, address, and location of your device
- personal interests and personal description stored on your profile
- images and videos stored on your account
- an acoustic model of voice characteristics
- phone numbers, addresses, and emails of your contacts
- information about your devices, such as specs and configuration
- which devices are linked to your smart assistant
- a record of interactions and requests made through your smart assistant, and
- internet provider and connectivity information, and more (Cohen, 2021).

While we acknowledge that collecting data is an integral part of improving a voice experience, people should know what type of personal voice data is being collected, why it is being collected, how it is being managed, and by whom.

For example, if voice biomarkers are collected for medical purposes, explain:

- what voice biomarkers are
- who is collecting them
- why they are being collected
- how they are collected and stored
- how long they will be stored
- which specific ones are being analyzed and for what purpose, and
- how the analysis will be interpreted.

People should also be informed when both Speech Emotion Recognition and sentiment analysis are run, as they both affect the impact of the data and what people are willing to share.

Disclose the purpose for every piece of personal voice data collected

Every piece of personal voice data needs to be accounted for and have a clear purpose that is disclosed in the terms and conditions provided to the user. If the purpose is for future research or product improvement, disclose it anyway.

People need to understand the benefits they get out of giving away sensitive voice data and why it is being used. These benefits should outweigh the harms, given how much can be collected and inferred from personal voice data.

Disclose all third parties involved with accessing, owning, processing, and analyzing personal voice data

List the names of the third parties, why they need access to the data, what they will use the data for, and how the data will be processed and analyzed. Provide links to each party's privacy policy and terms and conditions.

If a new party needs access to the personal voice data after consent is given, consent needs to be provided again.

Clarify when devices are listening vs. recording

A major concern and deterrent for the use of voice technology is the idea that devices are spying or listening in on private conversations at all times. Disclose when devices are listening, what they are listening for, what triggers a voice device to start recording, and exactly when recordings are started and stopped.

Explain how wake word detection works

A wake word is a word or phrase that activates a voice assistant and prompts a device to start recording. If a wake word is used, disclose how the wake word is being detected and when the device starts and stops recording when it detects the wake word.

Have a process in place for the withdrawal of consent for the use of personal voice data

Regardless of an opt-in or opt-out approach for consent, there should be an explicit and well-defined procedure to withdraw consent for the use of personal voice data at any time. There should also be a clear explanation of what happens to personal voice data after consent is withdrawn.

Withdrawing consent should be an easy and straightforward process with a maximum of two pieces of information:

1. The withdrawal request (mandatory)
2. The reason for withdrawing consent (optional)

Disclose how and where people can delete their personal voice data

The process for deleting personal voice data should be featured prominently and stated clearly in the terms and conditions, in accordance with local, state, federal, and international laws. The information should include how and where the voice data is deleted, how people can confirm that their voice data was deleted, and how the deletion is handled when third parties are involved. This is particularly important for [accidental data processing](#).

Write terms and conditions and privacy policies about personal voice data clearly so people of all literacy levels can understand them

Given the sensitive personal nature and complexity of the human voice and voice experiences, it is essential that all people clearly and easily understand what they are consenting to before they interact with voice technology.

Privacy Protection

Statistic

According to the Voice Consumer Index 2021 by the Open Voice Network in partnership with Vixen Labs and Delineate (2021), around 50% of people who use voice technology say they are concerned about privacy (p. 11). Privacy, trust, and safety are also major concerns and barriers for non-users (Open Voice Network, Vixen Labs, & Delineate, 2021, p. 12).

NOTE: As mentioned in the introduction, our guidelines focus on voice interactions in the physical world. We plan to release guidance once we conduct additional research and the legal framework for extended reality spaces becomes more well-defined.

The Universal Declaration of Human Rights by the United Nations (1948) recognizes privacy as a fundamental human right. Many nations worldwide also support and acknowledge privacy as a fundamental human right. Ensuring the privacy and protection of privacy of all people should be a priority from conception to the testing and launch phases, no matter where you are located.

Privacy protection for voice experiences is crucial for the following reasons:

- The biometric nature of voice and the amount of sensitive information that is shared with every voice interaction
- Accidental data processing happens
- Voice data that may contain sensitive information must be protected under existing laws, such as HIPAA and COPPA
- The lack of knowledge and understanding of all voice-specific security vulnerabilities or breaches that could occur.

It is particularly important to prevent situations such as one that happened with hackers remotely hijacking and taking control of a teakettle connected to the internet because of security vulnerabilities (American Bar Association, 2016). Security vulnerabilities with voice technology could lead to hackers or other third parties controlling voice-enabled devices, listening to and recording private audio conversations at any time, extracting personal voice data, and leaking or selling all kinds of personal information (American Bar Association, 2016). Any breaches in security could negatively impact trust in voice technology and cause serious liability issues.

There are also potentially undiscovered and uninvestigated voice-specific security vulnerabilities and breaches that exist because voice technology is developing and evolving, such as “attacks that hide malicious voice commands in nonsensical word sounds and in apparently unrelated utterances” (Bispham, 2020, p. 1). The harm that can come from both known and unknown attacks and the type and volume of data shared during voice interactions calls for special care and attention paid to people’s privacy. Part of our mission is to ensure everyone is aware of the issues that can occur and acts in accordance with common privacy best practices.

See the European Union General Data Protection Regulation (GDPR) for general guidance on privacy rights. For voice-specific privacy regulations, see the California Consumer Protection Act (CCPA) with follow-on amendments in the California Privacy Rights Act (CPRA), the Illinois Biometric Information Privacy Act, and the European Data Protection Board (EDPB) guidelines for virtual voice assistants. For guidance regarding children’s privacy, see the Children’s Online Privacy Protection Act.

Prioritize data security

Personal voice data should be stored safely and securely. If any sensitive voice data is processed, managed, or stored by third parties, links to all third-party privacy policies should be provided.

All organizations should have a procedure for dealing with personal voice data breaches. If a data breach occurs, users should be notified immediately. Detailed investigations and records should be kept regarding any breaches experienced by an organization in case users want information and explanations regarding the risks to which they have been exposed.

Best practices that build trust include:

- Obtaining consent for the use of personal voice data at the beginning of the contractual relationship

- Sharing clear and easily understandable information regarding the privacy policy for voice
- Including a separate section detailing what is unique about voice privacy if the services of the voice assistant are being provided as part of a broader set of services by an enterprise or organization
- Requesting explicit consent and disclosing a clear explanation when policy or software updates impact the privacy of users or data management procedures
- Informing all users of the privacy implications of engaging with third parties that handle personal voice data

Thoroughly vet third parties that are handling and processing personal voice data

Pay particular attention to people's privacy rights when engaging in third-party negotiations, especially when children could be involved. Make sure that the highest standards of privacy for voice experiences, especially the ones processing personal voice data, are guaranteed by both organizations.

Major tech companies typically send audio files to a mix of internal employees and contractors for review (Day et al., 2019). In the past, these companies have received backlash for not explicitly letting users know that humans were reviewing their conversations and ended up issuing rare apologies and making changes to their privacy settings and policies because of it (Nieva & Rubin, 2019). Although the purpose of human review is to train the algorithms and improve the experience, audio snippets can be tied back to individuals and privacy can be compromised if the audio is not properly anonymized.

Recordings from one tech company sent to contractors were tied to an account number, the user's first name, and the serial number of the device (Day et al., 2019). Contractors were able to hear highly personal information such as "lonely sounding people confessing intimate secrets and fears; a boy expressing a desire to rape; men hitting on Alexa like a crude version of Joaquin Phoenix in Her...Other contractors recall hearing kids share their home address and phone number, a man trying to order sex toys, a dinner party guest wondering aloud whether Amazon was snooping on them at that very instant" (Carr et al., 2019, para. 3).

We understand that human review of conversations is vital to improving voice experiences, but it cannot come at the expense of people's privacy. The harms that can occur from the leakage of personal voice data or the mishandling of information from third parties or unknown parties could be devastating. One company had to briefly stop all human review of audio from voice interactions after it was found that a contractor had been leaking parts of the audio (Newman, 2019).

Best practices for organizations and enterprises include ensuring all audio snippets are anonymized and cannot be linked back to the user when sent out for review, and giving people an option to opt out of the human review of any recorded voice interactions.

Minimize the amount of personal voice data being collected

For privacy and security purposes, do not collect personal voice data unless there is a specific reason for collecting it. If it is essential to collect personal voice data, only collect the minimum amount of voice data needed to fulfill a specific purpose, nothing more. Get to the root of why a particular piece of personal voice data is being collected, then figure out if it is worth collecting for that purpose. If more personal voice data is needed in the future, you can ask for consent again when that data is needed.

Use a voice anonymizer whenever possible

We understand there are challenges to using a voice anonymizer, particularly in the context of larger companies, as it dilutes the voice data necessary to improve voice experiences. However, be aware that it is an option and use it at your discretion.

Do not share sensitive personal voice data information with anyone that does not have permission to access it

This is particularly important when it comes to compliance with laws such as HIPAA and COPPA. People need to trust that their sensitive information will be protected and secured properly when they share it.

Delete personal voice data when it has fulfilled its purpose

Fully delete personal voice data in accordance with existing laws after it has fulfilled its purpose, which should be stated clearly in terms and conditions. Do not store personal voice data longer than necessary to avoid a greater risk of something negative occurring because you do.

Allow people to easily access and delete their personal voice data at any time

People should be able to request the deletion of part or all of their personal voice data at any time in accordance with local, state, federal, and international laws. This process should be timely, easy, and straightforward to understand. People should also receive confirmation that their data was deleted successfully.

The motivation for requesting deletion of sensitive voice data is not relevant to the actual deletion of data and should not be questioned.

Have a process in place for accidental data processing

Personal voice data can be recorded, stored, and processed without knowledge and consent if a device is accidentally activated by a word similar to its wake word (Schönherr et al., 2020). That can cause serious privacy issues as recordings are sent to the cloud and can potentially be processed and analyzed.

Lawsuits have been filed against Google, Amazon, and Apple, claiming that devices were activated and recorded private conversations even though the wake word wasn't used (Lerman, 2021). Those recordings were then used to target advertisements and allegedly given to third parties, which is a clear violation of privacy.

We recognize that accidental voice data processing is not always easy to catch, which is why a clear process for users and providers to continuously monitor, audit, and delete accidental voice data processing is important.

Inclusivity

Voice is a uniquely inclusive form of technology. Many people rely on voice technology to complete daily tasks such as keeping in touch with others, interacting with other Internet-enabled devices, and searching for information (Ammari et al., 2019). Take Kaden, for example.

Kaden is a boy with cerebral palsy who cannot walk or talk, is legally blind, and can use only the pinky finger on his left hand (Snow, 2019). Kaden uses a speaking device to communicate with his Amazon Echo Show by pressing buttons with his pinky to verbalize preselected words or sentences. His Echo Show was set up so that Kaden could ask for things such as the news or a joke and keep in touch with family, but Kaden has been using the technology beyond its intended use, calling his grandfather one evening through his Echo Show to ask if they could go for a car ride.

Voice assistance is also used for critical tasks such as calling for help in an emergency or connecting to a hotline if someone is in dire physical and mental distress (Graham, 2017). It is particularly useful in situations when hands or eyes are busy, such as cooking or driving. Language is becoming less of a barrier, as audio can potentially be translated from text to speech in any language.

Initiatives such as Translators Without Borders are leveraging speech technology to address the needs of refugees and fight disinformation regarding COVID-19 in Nigeria and Congo (TWB Communications, 2020). One of their long-term programs aims at creating a more balanced communication with populations in low-literacy contexts that have historically been marginalized.

By engaging with and learning from a diverse group of people, we can work toward including everyone in a voice-enabled future. Voice technology systems should also be assessed regularly to ensure they are fair, free of bias, and designed with the input and participation of a diverse range of users.

Provide multiple options for consenting to terms and conditions and privacy policies in accessible formats

Consent should be available in multiple modalities and languages for accessibility and ease. Voice alone will exclude some categories of people from expressing consent. Anyone should be able to legally give consent to use personal voice data and agree to the terms and conditions and privacy policies in accordance with local, state, federal, and international laws.

Assemble diverse teams to design and develop voice experiences

Ensuring diversity within teams developing voice experiences is the first step toward preventing societal harms, such as favoring one group arbitrarily over another (Li, 2020). Diverse and well-rounded teams can also minimize potential bias because of the range of experiences and perspectives that can be drawn from throughout the entire design process (Orduña, 2019). This is important, especially for speech-recognition tools.

There is already known bias against women in general and people with accents that use voice technology in cars (McMillan, 2011), and racial bias against Black individuals, as proven by a study that tested commercial speech-to-text tools used by five of the major tech companies (Koencke et al., 2020). Diverse teams will be able to spot and anticipate these types of scenarios and think more creatively, which will lead to more inclusivity in voice experiences and better scalability across enterprises and organizations (Shastri, 2020).

Proper training and education of these teams is crucial. The more knowledge the team has about how the technology works and the potential issues that can occur, the more they can mitigate those issues and educate others about them.

Test voice experiences with diverse groups of people at every step of the design process

Getting feedback from diverse groups can spark new ideas, uncover design issues, expose potential bias and discrimination, and create broader customer bases. It is particularly important to test wake words and invocation words. If people can't invoke your voice assistant, Skill, or Action, they will not be able to use it. Certain people might have a harder time saying specific letters or words, so make sure the words or names you use are as universally easy for people to say and remember as possible.

Also reach out to organizations, groups, and individuals that could potentially be excluded from the design due to challenges, disabilities, and underrepresentation, and include them in your testing groups. Google is currently testing Project Relate, an app that helps people with diverse speech patterns communicate with Google Assistant (Schwartz, 2021). This project will help

improve Google's speech recognition technology for people with ALS, Cerebral Palsy, Parkinson's disease, and more. The more variation you have, the more people will be able to interact with the design.

Gather large, diverse, and inclusive datasets for training voice AI models

Voice AI learns from data produced by humans, incorporating human bias into voice datasets and leading to algorithms that reflect human prejudices. Those biases need to be taken into account and mitigated when collecting and processing data.

The Intelligence Community (2020) says unintentional bias can "undermine analytic validity, harm individuals, or impact civil liberties such as freedom from undue government intrusion on speech, religion, travel, or privacy" and may be "introduced through the process of data collection, feature extraction, curating/labeling data, model selection and development, and even in user training" (sec. 5, para. 2).

Awareness of these biases can go a long way toward designing more inclusive voice products and services, but that alone is not enough. In addition to education, awareness, and documenting previously known biases, determine at which points humans should be involved in the decision-making process and the level of accountability required for each role through risk assessment (Intelligence Community, 2020). Source broad and diverse voice datasets and consider obtaining feedback from experienced ethicists to guide inclusive design.

Engage with and educate your communities

There is still a learning curve when it comes to voice interactions. It is easy for people to give up if a voice experience does not work for them. The more people know how to use voice technology and see the value, the more people can trust that it will work for them and make informed decisions about how to use it.

Engaging with your communities allows you to tailor features to human needs, learn what people want out of voice experiences, understand the pain points and how to overcome them, and figure out what might be missing. Listening to and implementing feedback will make the experience more accessible and beneficial to all. If one person says that your voice experience is not accessible, it is very likely that others have the same problem. Provide a way to gather feedback and create an active feedback loop that is continuously monitored to help incorporate suggestions.

Be aware of existing stereotypes when applying gender to voice assistants

Gender bias is particularly prevalent in voice assistants, often as a result of the makeup and biases of the team members creating them. In the tech industry, these teams tend to be predominantly male. Because a majority of personal assistant bots such as Siri, Alexa, and Cortana are perceived to be female, "it sends the signal that women are obliging, docile and

easy-to-please helpers, available at the touch of a button or with a blunt voice command like “Hey” or “OK” (UNESCO, 2019, p. 104).

This can negatively influence the way children who interact with voice assistants on a regular basis grow up to perceive and treat women. It also opens the door for what “Laura Bates (2014) calls ‘everyday sexism’: those small and seemingly harmless insinuations that women experience on a near-daily basis ... It is the kind of ‘harmless’ language that can contribute to women becoming the target of physical violence...” (as cited in Kennedy & Strengers, 2020, p. 61). This kind of abuse and harassment should be addressed through the voice assistant’s responses and used as an opportunity to raise awareness and educate people.

Other solutions for mitigating gender bias include:

- Offering a variety of voices and accents instead of a singular default voice whenever possible (Robison, 2020)
- Using color as a way to represent gender (Robison, 2020)
- Using gender-neutral or non-binary voices like Sam, a non-binary voice created by Accenture and CereProc (Behr, 2020)

Avoid using human names as default wake words

The way wake words are used as commands in voice technology can impact the way humans are treated in everyday life. People named Alexa report being subject to severe bullying as a direct result of the way people interact with and treat Amazon Alexa (Ard, 2021). Some have legally changed their names because of the harassment, but for one Alexa, it got to the point where she told her mother she wanted to kill herself (Ard, 2021). The best way to mitigate this unintentional consequence is to avoid using human names as default wake words or provide a way to set custom wake words on all devices.

Accountability

Given the vast amount of personal information that can be processed and analyzed from a voice, accountability is critical to building trust. Accountability means maintaining and following the highest possible ethical standards when managing sensitive voice data and holding people accountable for their actions when negative outcomes occur.

Be clear about the responsibilities of each stakeholder involved with the management of personal voice data from the start

Clearly defining the roles and responsibilities of each stakeholder will make it easier to hold people accountable if any negative outcomes occur as a direct result of the use of personal voice data.

Provide a way for people to give feedback if an unexpected outcome from the use of personal voice data occurs

Awareness is key to uncovering issues that might not get caught otherwise and preventing future issues from occurring. Given the sensitive nature of voice data, the more people can give feedback, the better it will be for improving future voice experiences.

Allow people to challenge outcomes that stem from the use of personal voice data

If someone disagrees with an outcome that occurred because of the use of personal voice data, they should be able to challenge the outcome. If it is determined that the outcome does not need to be rectified, at least people can understand how and why the outcome happened. If it is determined that the outcome should have been different, future outcomes of the same nature can be prevented.

Hold every party involved with voice experiences to the highest levels of integrity and ethical thinking standards

If a third party mishandles any personal voice data or does not comply with the standards and agreements set for compliance with rules and regulations, cut ties with the third party immediately and replace it with another party that will comply with the standards.

Hire or consult ethicists that can review and evaluate the impact of ethical decisions being made regarding personal voice data

Working with experts that can review and evaluate ethical decisions can help prevent potential issues and “enable the creation of well-defined policies that promote human rights, safety, and economic benefits” (IEEE, 2019, p. 206).

Conduct regular internal and external audits

Audits specifically related to personal voice data should be conducted to monitor ethical decisions and outcomes, measure success if key performance indicators have been set, and ensure that laws, standards, and guidelines are being followed.

Train all team members equally, thoroughly, and properly

Thorough and proper training for every employee on voice technology is essential, no matter their background or which part of the design process they are working on (Byrum, 2020). Lack of education cannot be an excuse for negative outcomes, especially regarding sensitive voice data.

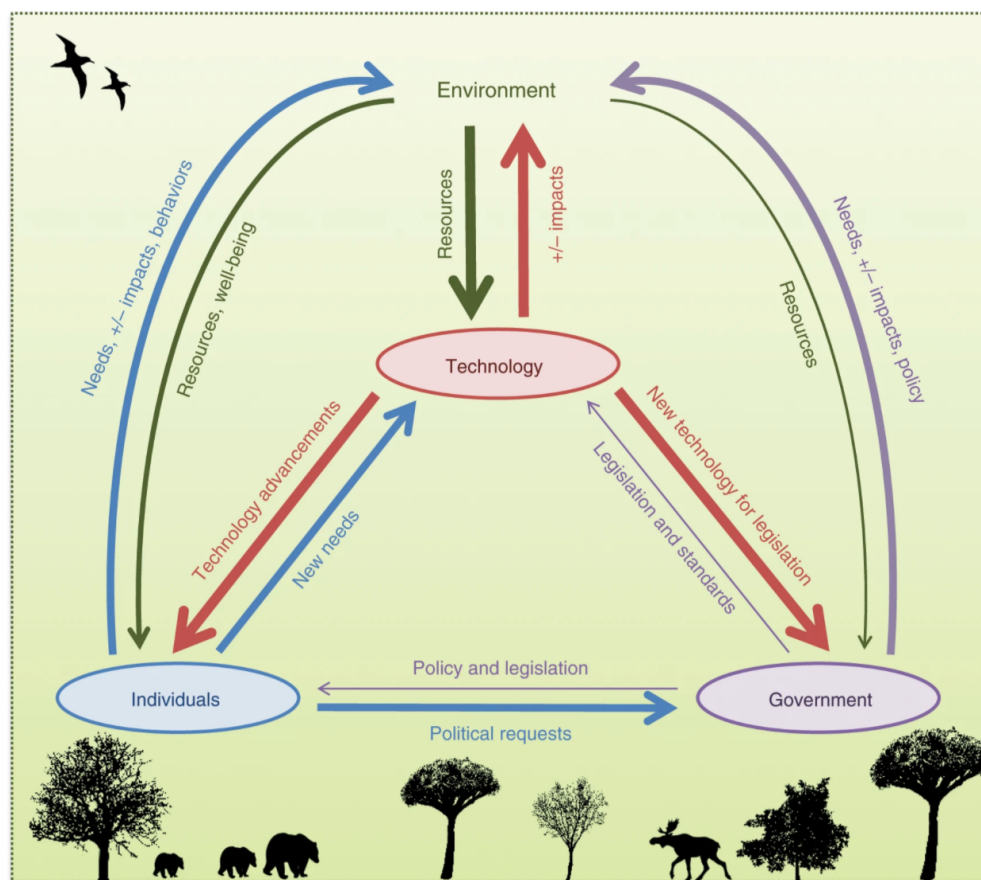
All team members need to “speak the same language” and “share a frame of reference” to successfully achieve their goals (Byrum, 2020, para. 10).

Sustainability

When discussing sustainability in voice technology, we are talking about generating net benefits, doing no harm, and protecting the environment. Voice technology can “leverage quality and better standards of life and protect people’s dignity, while maintaining cultural diversity and protecting the environment” (IEEE, 2019, p. 142). Ultimately, it is about a conscious consumption of technology and voice, or engaging with technology while being aware of the larger impacts on society (Wong, 2019).

Figure 2

Interaction of AI and society



Note. A schematic diagram of the interaction between AI and individuals, technology, government, and the environment. From “The role of artificial intelligence in achieving the Sustainable Development Goals,” by Vinuesa et al., 2020, *Nature Communications*, 11(233), <https://www.nature.com/articles/s41467-019-14108-y>. CC BY 4.0.

Generate net benefits for all stakeholders

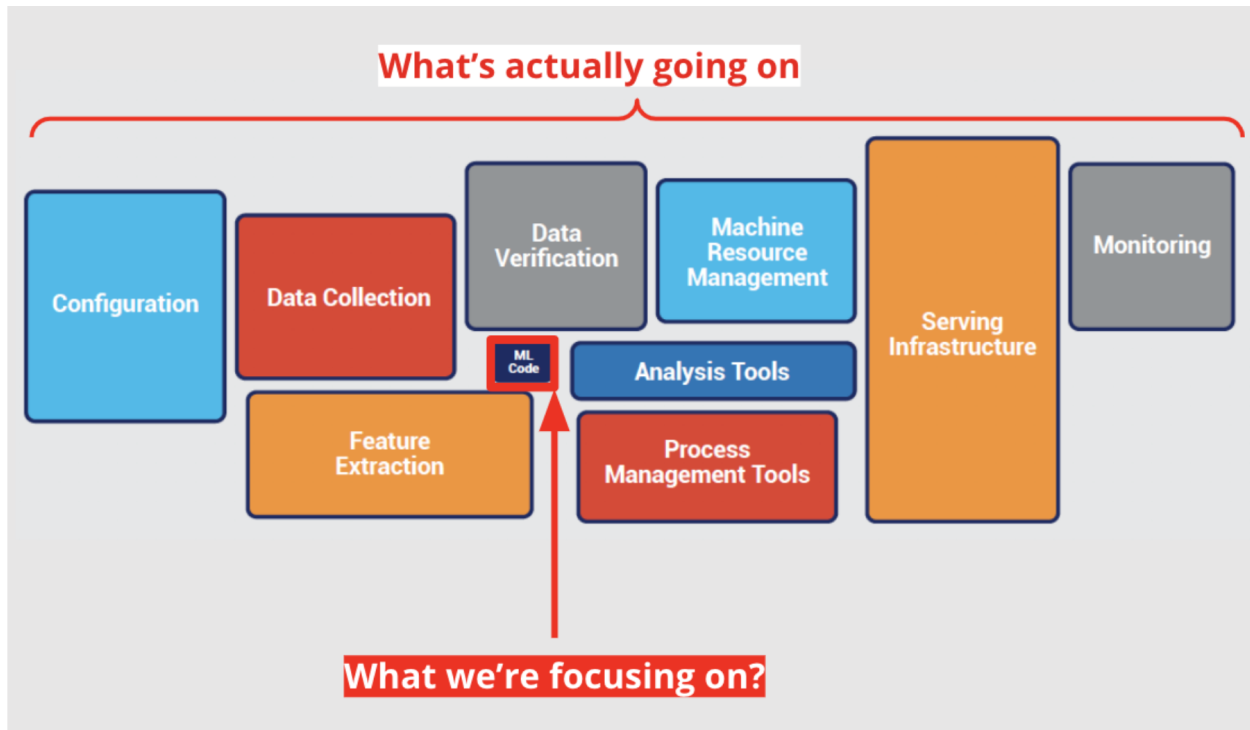
To provide net benefits for businesses, consumers, and the public sector, voice technology must have “universal respect for human rights and human dignity, the rule of law, justice, equality and nondiscrimination; ... respect for race, ethnicity, and cultural diversity; and ... equal opportunity permitting the full realization of human potential and contributing to shared prosperity” (United Nations, 2015, para. 8).

Be aware of the potential harms voice technology could inflict on the environment

“Put simply: each small moment of convenience—be it answering a question, turning on a light, or playing a song—requires a vast planetary network, fueled by the extraction of non-renewable materials, labor, and data” (Crawford & Joler, 2018, para. 8).

Figure 3

What we’re focusing on vs. what’s actually going on



Note. Diagram of the resources AI requires, pointing out that the majority of the focus is on machine learning and code. From “AI Ain’t Cheap,” by Bheemaiah & Esposito, 2021, Capgemini, <https://www.capgemini.com/2021/01/ai-aint-cheap/>. Copyright 2022 by Capgemini.

Voice AI training models “come with an economic trade-off, as they depend on the availability of exceptionally large computational resources that necessitate substantial energy consumption” (Bheemaiah & Esposito, 2021, para. 3). Even something as seemingly simple and innocent as “What’s the weather?” requires a robust network of resources to answer.

Do thorough research and be aware of the overall cost for designing and developing voice experiences. All stakeholders should be educated on the implications of using this technology. The people using voice technology have the power to shape how it is being used, but they have to have sufficient knowledge of how the technology works to make informed decisions about how to use it.

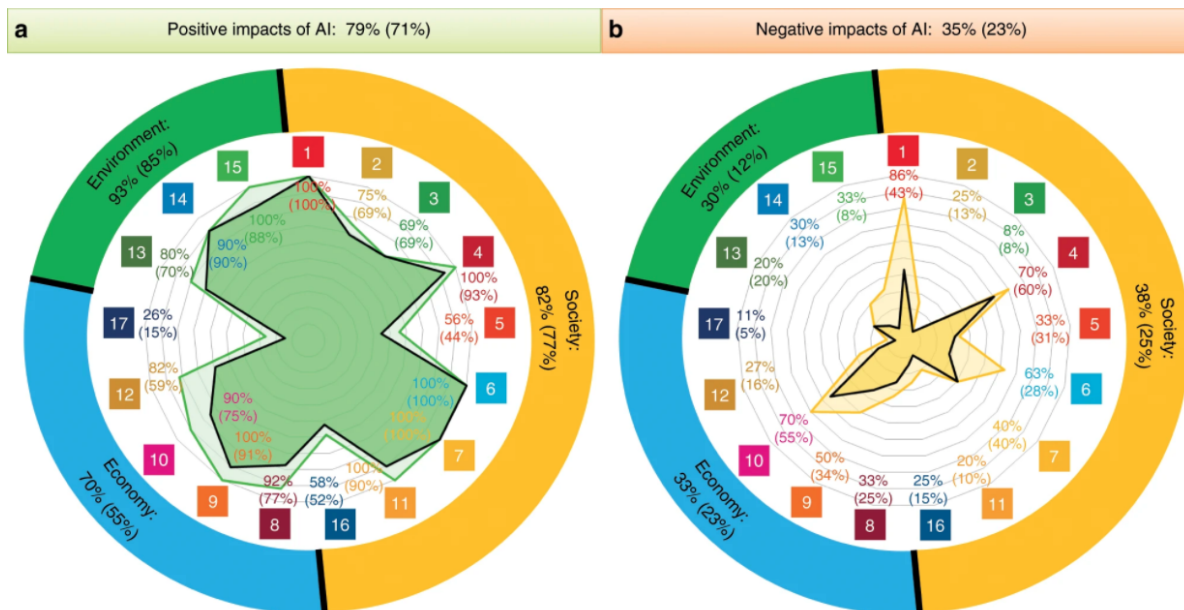
Continuously set goals and measure success

Create key performance indicators for sustainability that are continuously monitored and measured to see if the technology has the expected impact. An example of a goal could be driving down the energy consumption of call centers by increasing their efficiency and reducing the number of people who rely on them.

Measuring the success of these goals should indicate how to move toward a more sustainable future through voice technology.

Figure 4

Summary of positive and negative impact of AI on the various SDGs [Sustainable Development Goals]



Note. Percentage breakdown of potential positive and negative impacts of AI on the environment, society, and the economy. From “The role of artificial intelligence in achieving the Sustainable Development Goals,” by Vinuesa et al., 2020, *Nature Communications*, 11(233), <https://www.nature.com/articles/s41467-019-14108-y>. CC BY 4.0.

Make decisions on the type of voice technology used based on sustainability

Choose technology, implementation, and deployment based on your sustainability goals and values. One way to reduce energy consumption is to use solar power instead of electricity. For example, Haque et al. (2020) talk about an eco-friendly solar-powered and voice-controlled wheelchair built to help people with physical disabilities improve their self-mobility.

It might take a little innovation and creativity to find solutions that do not harm the environment, but there are options and resources out there.

Align with organizations that have the same sustainability values and goals

When collaborating with other organizations and enterprises, make sure they have similar sustainability goals. Partner with carbon-neutral companies when developing voice experiences or organizing awareness campaigns to educate people about the importance of sustainability.

Lead by example

Show people what is possible when sustainability is included in the voice technology decision-making process. If you can model it successfully, others will follow.

Call to Action

Start small

Implement small changes wherever you can at first, and then work your way up. It might be hard to see results right away, but results will show over time if you consistently implement changes.

Raise awareness about voice technology and continuously educate yourself and others

Voice technology is continuously growing and evolving. Keeping up with the latest news and updates, raising awareness about what voice technology can do for people, and educating people on the benefits and harms of voice technology are critical for its development and growth.

Continuously monitor and assess the state of voice technology and modify your decision-making process based on the most current state

Adaptability and flexibility as voice technology advances are going to be important for making ethical decisions that are in everyone's best interests. The Open Voice Network will continue to monitor and update our guidance, and we hope practitioners will continue to monitor these guidelines and provide feedback as well.

Continuously audit voice-specific decisions and processes and keep track of your progress over a period of time

Auditability of voice experiences over time will inform its creators on how they are doing, if they are progressing, and what changes could be needed to shape a trajectory of trustworthy voice experience design and development.

Get involved with the Open Voice Network

Provide feedback on the work that we are doing, share our work with others to amplify our message, attend our events, join our committees and workgroups, and reach out to collaborate with us.

About the Open Voice Network

The Open Voice Network (OVON) is a non-profit industry association dedicated to the development of standards for voice assistance transparency, consent, limited collection, and control of voice data that will make using voice technology worthy of user trust. In any reality, virtual or otherwise, we believe personal privacy should be respected as the default. The Open Voice Network operates as an open-source community within The Linux Foundation. It is independently funded and governed with participation from more than 120 voice practitioners and enterprise leaders from 12 countries.

The Open Voice Network community's work is open source. We seek inclusive input and like to share our insights. At present, our work is focused in four areas:

- **Interoperability**, defined as the ability for conversational agents to share dialogs (and accompanying context, control, and privacy),
- **Destination registration and management**, the ability of users to confidently find a destination of choice through specific requests, and for the providers of goods and services to register a verbal "brand"—similar to the Domain Name System (DNS) of the internet;
- **Privacy**, with voice-specific guidance for both the protection of individual user data and that of commercial users; and
- **Security**, with a focus on voice-specific threats and harms.

Please see our papers in 2022 and support the Open Voice Network by visiting openvoicenetwork.org.

About The Linux Foundation

Founded in 2000, The Linux Foundation is supported by more than 1,000 members and is the world's leading home for collaboration on open-source software, open standards, open data, and open hardware. Linux Foundation's projects are critical to the world's infrastructure including Linux, Kubernetes, Node.js, and more. The Linux Foundation's methodology focuses on leveraging best practices and addressing the needs of contributors, users, and solution providers to create sustainable models for open collaboration. For more information, please visit us at linuxfoundation.org.

Acknowledgements

Authored by the Open Voice Network with special thanks to Ethical Use Task Force contributors, Audrey Amsellem, Audrey Arbeeney, Dr. Maria Aretoulaki, Valeria Barbero, Emily Bazhaf (document draft), Leigh Clark, Oita Coleman, David Devoto, Brooke Hawkins, John T. LaGrand, Janice Mandel, Michael Novak, Sara Oliver, Leslie Pound, Emerson Sklar, Emily Troutner, and Nigel Write; edited by Janice Mandel.

Reference List

- American Bar Association. (2016, May 20). "Who am I talking to?" - The regulation of voice data collected by connected consumer products. Business Law Today. https://www.americanbar.org/groups/business_law/publications/blt/2016/05/06_black/
- Ammari, T., Kaye, J., Tsai, J. Y., & Bentley, F. (2019). Music, search, & IoT: How people (really) use voice assistants. *ACM Transactions on Computer-Human Interaction*, 26(3), 17-1. <https://doi.org/10.1145/3311956>
- Ard, A. J. (2021, December 3). Amazon, can we have our name back?. *Washington Post*. https://www.washingtonpost.com/technology/interactive/2021/people-named-alexa-name-change-amazon/?itid=sf_technology_article_list
- Barrett, B. (2019, August 2). Hey, Apple! Opt out is useless. Let people opt in. *Wired*. <https://www.wired.com/story/hey-apple-opt-out-is-useless/>

- Baugh, J. (2002). Linguistic Profiling. In S. Makoni, G. Smitherman, A. F. Ball, & A.K. Spears (Eds.), *Black linguistics: Language, society, and politics in Africa and the Americas* (1st ed., pp. 155-168). Routledge. <https://doi.org/10.4324/9780203986615>
- Behr, M. (2020, December 17). *Meet Sam: The non-binary voice for a digital assistant*. Digit. <https://www.digit.fyi/meet-sam-the-non-binary-voice-for-a-digital-assistant/>
- Bheemaiah, K., & Esposito, M. (2021). *AI ain't cheap*. Capgemini. <https://www.capgemini.com/2021/01/ai-aint-cheap/>
- Bispham, M. K. (2020). *The security of human-computer interaction by speech* [PhD thesis, University of Oxford]. <https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.820745>
- Burns, V., (2021, April 2). *How voice recognition is helping HR save time and hire smarter*. Unleash. <https://www.unleash.ai/how-voice-recognition-is-helping-hr-save-time-and-hire-smarter/>
- Byrum, J. (2020, May 18). *Build a diverse team to solve the AI riddle*. MIT Sloan Management Review. <https://sloanreview.mit.edu/article/build-a-diverse-team-to-solve-the-ai-riddle/>
- Carr, A., Day, M., Frier, S., & Gurman, M. (2019, December 11). Silicon Valley is listening to your most intimate moments. Bloomberg Businessweek. <https://www.bloomberg.com/news/features/2019-12-11/silicon-valley-got-millions-to-let-siri-and-alexa-listen-in>
- Chen, A. (2019, March 14). *Why companies want to mine the secrets in your voice*. The Verge. <https://www.theverge.com/2019/3/14/18264458/voice-technology-speech-analysis-mental-health-risk-privacy>
- Chin, C., & Robison, M. (2020, November 23). *How AI bots and voice assistants reinforce gender bias*. Brookings Institute. <https://www.brookings.edu/research/how-ai-bots-and-voice-assistants-reinforce-gender-bias/>
- Clark, L., Cowan, B. R., Roper, A., Lindsay, S., & Sheers, O. (2020). Speech diversity and speech interfaces: Considering an inclusive future through stammering. *Proceedings of the 2nd Conference on Conversational User Interfaces (CUI '20)*. Association for Computing Machinery, New York, NY, USA, Article 24, 1-3. <https://doi.org/10.1145/3405755.3406139>
- Haque, A. B., Shurid, S., Juha, A. T., Sadique, M. S., & Asaduzzaman, A. S. M. (2020, February 2-5). *A novel design of gesture and voice controlled solar-powered smart wheel chair with obstacle detection* [Paper, pp. 23-28]. IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), Doha, Qatar. <https://doi.org/10.1109/ICIoT48696.2020.9089652>
- Claypoole, T. F. (2021). Voice analysis complicates personal privacy. *National Law Review*, XI(271). <https://www.natlawreview.com/article/voice-analysis-complicates-personal-privacy>
- Cohen, J. (2021, October 22). *Amazon's Alexa collects more of your data than any other smart*

- assistant*. PC Mag. <https://www.pcmag.com/news/amazons-alexa-collects-more-of-your-data-than-any-other-smart-assistant>
- Crawford, K., & Joler, V. (2018, September 7). Anatomy of an AI System: The Amazon Echo as an anatomical map of human labor, data, and planetary resources. *AI Now Institute and Share Lab*. <https://anatomyof.ai/>
- Day, M., Turner, G., Drozdiak, N. (2019, April 10). *Amazon workers are listening to what you tell Alexa*. Bloomberg. <https://www.bloomberg.com/news/articles/2019-04-10/is-anyone-listening-to-you-on-alexa-a-global-team-reviews-audio>
- Despotovic, V., Ismael, M., Cornil, M., Mc Call, R., & Fagherazzi, G., (2021). Detection of COVID-19 from voice, cough, and breathing patterns. *Computers in Biology and Medicine*, 138, 1-9. <https://doi.org/10.1016/j.combiomed.2021.104944>
- Fagherazzi, G., Fischer, A., Ismael, M., & Despotovic, V., (2021). Voice for health: The use of vocal biomarkers from research to clinical practice. *Digital Biomarkers*, 5(1), 78-88. <https://doi.org/10.1159/000515346>
- Gershgorn, D. (2019, September 26). *Why it's so hard for Amazon Alexa to really explain itself*. Medium. <https://onezero.medium.com/why-its-so-hard-for-amazon-alexa-to-really-explain-itself-9d5ede67a68>
- Graham, J. (2017, July 19). 'Alexa, call 911' won't work. Here's what will. *USA Today*. <https://www.usatoday.com/story/tech/talkingtech/2017/07/19/alexa-cant-dial-911-but-google-alexa-and-siri-can-get-you-help/486075001/>
- Haque, A. B., Shurid, S., Juha, A. T., Sadique, M. S., & Asaduzzaman, A. S. M. (2020, February 2-5). *A novel design of gesture and voice controlled solar-powered smart wheel chair with obstacle detection* [Paper, pp. 23-28]. IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), Doha, Qatar. <https://doi.org/10.1109/ICIoT48696.2020.9089652>
- IEEE. (2019). *Ethically Aligned Design* (1st ed.). <https://standards.ieee.org/wp-content/uploads/import/documents/other/ead1e.pdf>
- Ingale, A. B., & Chaudhari, D. S. (2012). Speech Emotion Recognition. *International Journal of Soft Computing and Engineering (IJSCE)*, 2(1). 235-238.
- Intelligence Community. (2020). *Artificial intelligence ethics framework for the intelligence community*. <https://www.intelligence.gov/artificial-intelligence-ethics-framework-for-the-intelligence-community#Review>
- Kennedy, J., & Strangers, Y. (2020). *The smart wife: Why Siri, Alexa, and other smart home devices need a feminist reboot*. MIT Press.
- Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., Toups, C., Rickford, J. R.,

- Jurafsky, D., & Goel, S. (2020). Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences (PNAS)*, 117(14), 7684-7689. <https://doi.org/10.1073/pnas.1915768117>
- Laver, J. (1980). *The phonetic description of voice quality: Cambridge studies in linguistics*. Cambridge University Press.
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50-80. https://doi.org/10.1518/hfes.46.1.50_30392
- Lerman, R. (2021, September 2). Lawsuits say Siri and Google are listening, even when they're not supposed to. *Washington Post*. <https://www.washingtonpost.com/technology/2021/09/02/apple-siri-lawsuit-privacy/>
- Li, M. (2020, October 26). *To build less-biased AI, hire a more diverse team*. Harvard Business Review. <https://hbr.org/2020/10/to-build-less-biased-ai-hire-a-more-diverse-team>
- Maghilnan, S., & Kumar, M. R. (2017, June 23-24). *Sentiment analysis on speaker specific speech data* [Paper, pp. 1-5]. 2017 International Conference on Intelligent Computing and Control (I2C2), Coimbatore, India. <https://doi.org/10.1109/I2C2.2017.8321795>
- Mayeux, R. (2004). Biomarkers: Potential uses and limitations. *NeuroRX: the Journal of the American Society for Experimental NeuroTherapeutics*, 1(2), 182-188. <https://doi.org/10.1602/neurorx.1.2.182>
- McMillan, G. (2011, June 1). *It's not you, it's it: Voice recognition doesn't recognize women*. Time. <https://techland.time.com/2011/06/01/its-not-you-its-it-voice-recognition-doesnt-recognize-women/>
- Newman, L.H. (2019, September 23). *Google tightens its voice assistant rules amid privacy backlash*. Wired. <https://www.wired.com/story/google-assistant-human-transcription-privacy/>
- Nieva, R., & Rubin, B. F. (2019, September 25). *Amazon Alexa adds new commands to tamp down privacy concerns*. CNET. <https://www.cnet.com/home/smart-home/amazon-alexa-adds-new-commands-to-tamp-down-privacy-concerns/>
- Open Voice Network, Vixen Labs, & Delineate (2021). *Voice Consumer Index 2021*. <https://vixenlabs.co/voice-consumer-index>
- Orduña, N. (2019, July 16). *AI-driven companies need to be more diverse. Here's why*. World Economic Forum. <https://www.weforum.org/agenda/2019/07/ai-driven-companies-need-to-be-more-diverse-here-s-why/>
- Rice, P. (2006, February 2). Linguistic profiling: The sound of your voice may determine whether you get that apartment or not. *The Source, Washington University in St. Louis*. <https://source.wustl.edu/2006/02/linguistic-profiling-the-sound-of-your-voice-may-determine-if-you-get-that-apartment-or-not/>

- Robison, M. (2020, December 9). *Voice assistants have a gender bias problem. What can we do about it?*. Brookings Institute. <https://www.brookings.edu/blog/techtank/2020/12/09/voice-assistants-have-a-gender-bias-problem-what-can-we-do-about-it/>
- Schönherr, L., Golla, M., Eisenhofer, T., Wiele, J., Kolossa, D., & Holz, T. (2020, August 2). *"Unacceptable, where is my privacy?" Exploring accidental triggers of smart speakers*. arXiv. <https://arxiv.org/abs/2008.00508>
- Schwartz, E. H. (2021, November 9). *Google tests voice assistant for people with speech impairments*. Voicebot.ai. <https://voicebot.ai/2021/11/09/google-tests-voice-assistant-for-people-with-speech-impairments/>
- Shastri, A. (2020, July 1). *Diverse teams build better AI. Here's why*. Forbes. <https://www.forbes.com/sites/arunshastri/2020/07/01/diverse-teams-build-better-ai-heres-why/?sh=662f3e4177b3>
- Snow, J. (2019, January 30). *How people with disabilities are using AI to improve their lives*. PBS. <https://www.pbs.org/wgbh/nova/article/people-with-disabilities-use-ai-to-improve-their-lives/>
- TWB Communications. (2020, September 1). *TWB and KoBo Inc develop speech recognition technology to capture voices of speakers of marginalized languages*. Translators Without Borders. <https://translatorswithoutborders.org/twb-and-kobo-inc-develop-speech-recognition-technology-to-capture-voices-of-speakers-of-marginalized-languages/>
- UNESCO. (2019). *I'd blush if I could: Closing gender divides in digital skills through education*. <https://en.unesco.org/ld-blush-if-i-could>
- UNI Global Union. (2017). *Top 10 principles for ethical artificial intelligence*. http://www.thefutureworldofwork.org/media/35420/uni_ethical_ai.pdf
- United Nations. (1948). *Universal Declaration of Human Rights*. <https://www.un.org/sites/un2.un.org/files/udhr.pdf>
- United Nations General Assembly. (2015). *Transforming our world: The 2030 agenda for sustainable development*. <https://sdgs.un.org/2030agenda>
- Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Nerini, F. F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, 11(233). <https://www.nature.com/articles/s41467-019-14108-y>
- Wanucha, G. (2019, April). Talk about a revolution: The future of voice biomarkers in the neurology clinic. *Dimensions: The Magazine of the UW Alzheimer's Disease Research Center, Spring/Summer 2019*, 6-8. <https://indd.adobe.com/view/80228d6e-dfd7-4fd1-bfe7-57316f4e2a6b>
- West, D. M. (2018, September 13). *The role of corporations in addressing AI's ethical dilemmas*.

Brookings Institute. <https://www.brookings.edu/research/how-to-address-ai-ethical-dilemmas/>

Wong, K. (2019, October 1). How to be a more conscious consumer, even if you're on a budget. *New York Times*. <https://www.nytimes.com/2019/10/01/smarter-living/sustainable-shopping-conscious-consumer.html>

Wroge, T. J., Özkanca, Y., Demiroglu, C., Si, D., Atkins, D. C., & Ghomi, R. H. (2018, December 1). *Parkinson's disease diagnosis using machine learning and voice* [Paper, pp. 1-7]. IEEE Signal Processing in Medicine and Biology Symposium, Philadelphia, Pennsylvania, USA. <https://doi.org/10.1109/SPMB.2018.8615607>

Licensing and Attribution

Happy to share our work. Before you cite it, here's what you need to know:

This Open Voice Network white paper has been developed within the Vocabulary and Definition Project of the Open Voice Network, which has been established as Vocabulary and Definition Project a Series of LF Projects, LLC (the "Project"). LF Projects, LLC ("LF Projects") is a Delaware series limited liability company.

LF Projects hold title to all trade or service marks used by the Project ("Project Trademarks"), whether based on common law or registered rights. Project Trademarks will be transferred and assigned to LF Projects to hold on behalf of the Project. Any use of any Project Trademarks by Collaborators in the Project will be in accordance with the license from LF Projects and inure to the benefit of LF Projects.

Specification and documentation has been received and made available by the Project under the Creative Commons Attribution 4.0 International License (available at <http://creativecommons.org/licenses/by/4.0/>). To the extent a contribution includes or consists of data, any rights in such data shall be made available under the CDLA-Permissive 1.0 License.